DOCUMENT RESUME

ED 379 328 TM 022 704

AUTHOR Manatt, Richard P.
TITLE A Total Systems Approach to Performance Evaluation: How the School Improvement Model (SIM) Uses Evaluation To Improve Teaching and Learning.
PUB DATE Jul 94
NOTE 15p.; Paper presented at the Annual National Evaluation Institute of the Center for Research on Educational Accountability and Teacher Evaluation (3rd, Gatlinburg, TN, July 10-15, 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Uses in Education; *Curriculum Evaluation; Educational Assessment; *Educational Improvement; Educational Objectives; Elementary Secondary Education; Evaluation Methods; *Evaluation Utilization; Feedback; Learning; Models; Student Evaluation; Supervision; *Systems Approach; *Teacher Evaluation; Teaching Methods
IDENTIFIERS *School Improvement Model

ABSTRACT

The School Improvement Model of teacher performance evaluation is described, and its development is traced. Original work in the late 1960s in Naperville (Illinois) and subsequent work by Iowa State University and several cooperating school districts led to the development of a model of teacher evaluation that included multiple appraisers, student feedback about teachers, formative and summative steps, a modified clinical supervision cycle, and a written agreement for improved performance in the next cycle. The model has been developed and refined in actual practice in a series of school districts and consortia since 1978. It features both curriculum assessment and feedback. Goals and standards are established, and teachers are enabled to use both in teaching. Checks on student performance are used not only for grading and evaluation of students but to describe individual learning difficulties through feedback and to prescribe specific remediation and reteaching procedures. Current work with a microcomputer-based system is proving demanding but worthwhile in pursuit of the Goals 2000 objectives. One table and four figures illustrate the discussion. (Contains 19 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 379 328

Paper presented at the

CREATE National Evaluation Institute

July 10-15, 1994

Gatlinburg, Tennessee

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
RICHARD P. MANATT
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Center for Research on Educational Accountability
and Teacher Evaluation (CREATE)

The Evaluation Center
Western Michigan University
Kalamazoo, MI 49008

A Total Systems Approach to Performance Evaluation: How the School Improvement Model (SIM) Uses Evaluation To Improve Teaching and Learning.

The Institute was supported in part by the Office of Educational Research and Improvement, U.S. Department of Education, (Grant No. R117Q00047). The opinions expressed are those of the authors, and no official support of these positions by the U.S. Department of Education is intended or should be inferred.

BEST COPY AVAILABLE

TM022704

TM

A Total Systems Approach to Performance Evaluation:

How the School Improvement Model (SIM) Uses Evaluation to Improve Teaching and Learning

an address by
Richard P. Manatt, Director
School Improvement Model
and
Chair, Educational Administration Programs
Iowa State University
515/294-5521

Third Annual National Evaluation Institute
CREATE
Gatlinburg, Tennessee
Monday, July 11, 1994

During the late 1960s, one of my former students, Dr. John Fields, became the superintendent of the Naperville, Illinois public schools. At that time, the elementary and secondary units were separate; he had two boards and two districts, but they were coterminous. Naperville was to experience explosive growth as the Argonne Laboratory, Standard Oil's Credit Card Center, and other clean industry located in the vicinity brought thousands of families to the area.

Superintendent Fields called my office with an interesting proposal. He offered to use his district as a large experiment to create a valid, reliable, and legally discriminating performance appraisal system for teachers. He made this challenge, "I'll take care of the union, Dick, if you'll take care of the research."

For the Naperville project, we used mass authorship to establish the social validity of the performance criteria, *i.e.*, we surveyed the entire faculty regarding the appropriateness of the proposed criteria. Next, we established the measure of teacher effectiveness to be rather rudimentary criterion-referenced tests. Not surprisingly, the faculty was not very happy about the notion of teacher ratings coupled with student outcomes!

Because we had this splendid opportunity to "data snoop," we played with the predictive power of five different data sources, *viz.*, self-ratings, one administrator, three administrators, three peers, and all of the students in the second class of the day taught by the teacher. (Hidlebaugh, 1973) Our rationale as explained to the teachers has been expressed well in present times by Don Medley:

> There is one defensible and important use for measures of the effectiveness of the individual teachers, that is as a basis for validating ratings of teacher performance. Validation of ratings (or any process-based teacher evaluation procedure) depends on evidence that students taught by teachers who receive higher performance ratings learn more, on the average, that students taught by teachers who receive lower performance ratings. This requires measures of effectiveness of individual teachers. (Medley, 1992)

The sources of data were similar to the five keys suggested by Duke and Stiggens 15 years later (Duke and Stiggens, 1986). They varied considerably in discrimination power (See Figure 1). This project identified 24 performance criteria (teacher behaviors).

Figure 1. Discrimination power of ratings from five appraiser types (Source: Manatt, Palmer, and Hidlebaugh, 1976).

Appraiser	Rank in order of Discrimination
Students	1 = High
Three Peers	2
Three Administrators	3
One Administrator	4
Self-Appraisal	5

The criteria set has been expanded and refined over the past 15 years and they always make sense to the stakeholders' groups that we used to direct the development of performance evaluation instruments (Stow and Sweeney, 1981; Manatt and Stow, 1984; Manatt and

Stow, 1986; Manatt and Daniels, 1991). Interestingly enough, these criteria correspond closely to the duties-based checklist proposed by Michael Scriven (Figure 2).

Figure 2. A comparison of Scriven's duties-based criteria and those identified by SIM (Sources: Scriven, 1988; Manatt and Stow, 1984).

Scriven	SIM
1. Know the subject matter	Crit. 7. Displays knowledge of subject matter
2. Design instruction	Crit. 2. Designs lessons
3. Select and create materials	Crit. 8. Selects/creates learning content
4. Construct tests	Crit. 6. Evaluates student work
5. Grade or mark students' performance	Crit. 6. Evaluates student work
6. Provide information to students about their achievements	Crit. 5. Provides students with special feedback
7. Provide information to administration	Crit. 21. Provides accurate data to school
8. Provide information to parents, guardians, and authorities	Crit. 21. Provides accurate data to parents
9. Uses resources	Crit. 12. Makes use of time, materials, resources
10. Communicates effectively	Crit. 4. Teacher communicates effectively
11. Manage the classroom	Crit. 13. Demonstrates personal organization
12. Engage in self-evaluation and development	Crit. 22. Keeps practices current
13. Render service to the profession	Crit. 24. Responsibilities beyond the classroom
14. Acquire and use knowledge of the school and community	Crit. 24. Responsibilities beyond the classroom

Once we had an evaluation instrument for the summative, end-of-the year report, the stakeholders' group turned its attention to procedures to be followed by the multiple appraisers that Dr. Fields and his school board wanted, a combination of what CREATE calls a hybrid model was chosen (Scriven, Wheeler, and Haertel, 1991).

All of us agreed that the student ratings of teachers were to remain an integral part of the model. Next, we added George Redfern's Job Improvement Targets to the end of the cycle (a modified MBO model). Finally, clinical supervision was chosen as the classroom observation model. We used the writing and research reports of Cogan, Goldhammer, and Anderson. Goldhammer had been a doctoral student who moved with Cogan from Harvard to the University of Pittsburgh. Goldhammer's dissertation centered on the analysis of data he collected while working in the Harvard program of clinical supervision for the Master of Arts in Teaching students. Goldhammer died before finishing his book, *Clinical Methods for the Supervision of Teachers* (1969). The final editing was done by Robert Anderson, who was then on the faculty at Harvard.

The Iowa State University contribution to the Naperville teacher performance evaluation model stopped with the training of appraisers to use the five-step model that Cogan built upon the three-step procedure for encouraging self-analysis suggested by Anderson, Barr, and Bush as early as 1925. In addition to the steps of "preteaching

conference," a "classroom visit," and a "follow-up conference," Cogan had added a labeling and analyzing step and a postmortem recap after the process ended.

The Naperville districts were well satisfied with their new teacher evaluation system. The careful, three-year approach of (1) develop the system, (2) pilot-test the system for one year, and (3) refine and institutionalize the system, became our *modus operandi* for the projects of the next years. Naperville immediately requested the development of an administrator performance evaluation system which consumed another three years.

Refinement Through Repetition

Through word of mouth, presentations at state and national association meetings, stakeholders" approach to creation of new performance evaluation systems spread. Dr. Charles Joss, superintendent of West Des Moines Community Schools, called and requested teacher and administrator performance evaluation. Dr. Bill Anderson wanted good teacher performance evaluation for the Des Moines Public Schools, which added the subsystem of a program of intensive assistance for marginal teachers. Mount Prospect, Illinois was next, followed by Pasadena, Garvey School District (Rosemead and El Monte), and Novato, all in California. Each repetition gave us a chance to sharpen our methods, get a better handle on costs, and to refine our model.

Ever since our first modest effort to study teacher performance evaluation criteria 20 years ago, we have maintained a consistent set of guiding principles (philosophic premises if you will) that provide direction to working with a client district First, what we do is a process not a product. Each model, because it is planned for, operated by, and controlled by the teachers and administrators of that school organization, is unique. Therefore, the components will vary but the philosophy will not. A few of our slogans used with stakeholders' committees will illustrate.

- "You can't improve a school without changing behaviors of the employees."
- "Effective classrooms are nested in effective schools."
- "Teachers are the solution, not the problem."
- "Quality control is a management decision."
- "Start with the board and administrator performance because good bosses set good examples."
- "People don't do what you expect, they do what you inspect."
- "What gets measured gets accomplished."
- If you want it taught, test for it!"

Now that may sound like a heavy-handed accountability press that would lead to draconian measures for summative evaluation which would be repugnant to teachers. Quite the contrary; all development is planned by a stakeholders' committee on which teachers will be the dominant group.

The stakeholders' committee is appointed by the superintendent and, in the initial charge to the group, it is made clear that the task is very important and that each member of the group was appointed for his/her special knowledge and skills. Stakeholders are selected to represent teachers, administrators, parents, students, and board members. Once more than half of the stakeholders are to be teachers, and, typically, the teachers are selected by the leaders of their association or union.,

Stakeholders are made to understand from the outset that they are an *ad hoc* group serving at the pleasure of the board of education and that their assignment is to "decide to recommend"; the board must make the final decisions on policy regarding performance appraisal. Stakeholders committees vary in size from 15-25 members, depending on the size of the school organization.

At first glance it might appear that the preponderance of teacher members would result in a water-down performance appraisal system. That never happens for several

reasons. Most important, teachers have higher standards than those of other stakeholders. Teachers invariably press for more data-gathering, more analyses, and better records and conferences. Teachers and administrators get caught up in the challenge of creating both an administrator and teacher performance e valuation system.

Eventually, we posed six questions to guide the work of stakeholders.

1. What are your purposes for performance evaluation?
2. How shall we determine the criteria?
3. How high shall we set the standards?
4. How shall be monitor and report performance?
5. How do we help evaluatees improve after we have their performance profiles?
6. How much training for evaluators and evaluatees is necessary for success?

Stakeholders are expected to serve for at least three years. In Year One, a custom-tailored performance evaluation system is planned by the stakeholders. All professional positions are included in the plan. During the second year, and after approval by the board of education (and the collective bargaining process where required by law), each principal and a couple of teachers in each building test the proposed system. After careful analysis of the pilot test, the system is refined and resubmitted to the collective bargaining process if required. Once approved, this new system is taught to all educational personnel. The total-systems approach to performance evaluation is used during the third year and specific inservice activities are added for appropriate personnel. Again, approval from the collective bargaining process may be required.

Good organizations don't just measure your competence, they teach you to be more competent. Thus, building upon contemporary research bases of school effectiveness and classroom effectiveness, the SIM activities include:

1. evaluating and improving the performance of all administrators (including the superintendent and the board of education),
2. evaluating and improving the performance of teacher,
3. designing and implementing a staff development and training component to (a) operate the new monitoring system successfully and to (b) change administrative and teaching behaviors to maximize learning for students, and
4. renewing curriculum and measuring how much students learn.

Generally speaking, the stakeholders identify seven functions for an articulated performance evaluation system:

1. To improve teaching and administration (identify ways to change teaching systems, environments, behaviors, change management systems, climate behaviors.)
2. To supply information that will lead to modification of assignments such as placement in other positions, promotions, terminations.
3. To protect students from incompetent teachers and teachers from unprofessional administrators.
4. To reward superior performance.
5. To validate the school organization's teacher/administrator selection process.
6. To provide a basis for teachers' and administrators' planning and professional development.
7. To raise student achievement.

By the late 1970s, the processes were working smoothly and the total-systems approach repeatedly resulted in a hybrid model of evaluation which included multiple appraisers, student feedback of teachers, formative and summative steps, a modified clinical supervision cycle, and a written agreement (either a Job Improvement Target, a Professional Improvement Commitment, or a Professional Growth Plan) for improved performance in the next cycle.

The Original SIM Project

By 1977-78 the performance appraisal research at Iowa State University had become established. In Spring 1978 the Northwest Area Foundation of St. Paul provided a major grant to attempt a huge experiment with several school organizations to improve instruction and to provide an accountability model (Manatt and Stow, 1986). The organizations included Spirit Lake, Iowa; Breck Independent School, Minneapolis Public Schools, Northfield Public Schools, and Edina Public Schools, all in Minnesota.

Conclusions

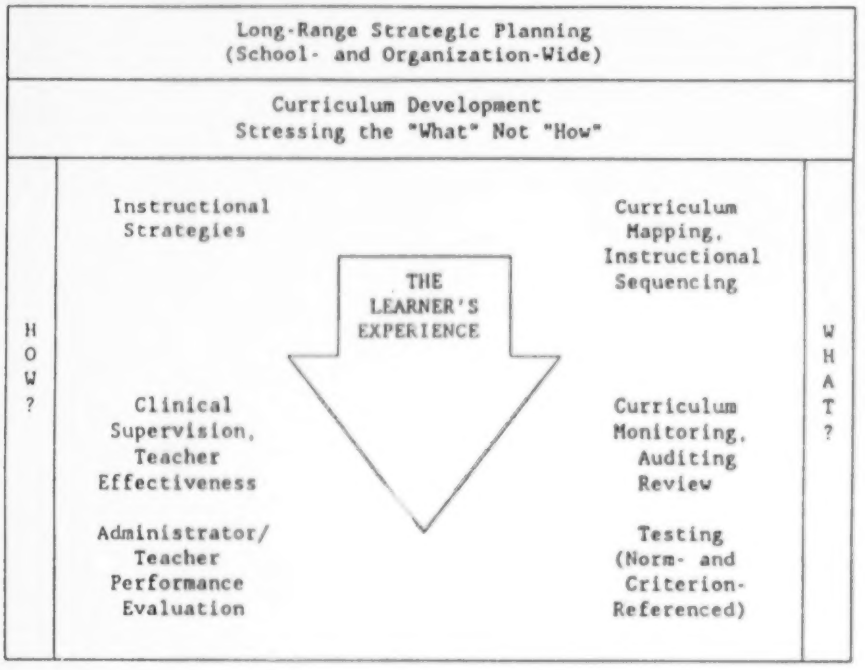
The paramount finding of this study was that principals can accurately evaluate the performance of teachers. When principals were given extensive training and when the limitations of earlier studies (Medley & Coker, 1987) regarding instrumentation and methodology are overcome, they are good judges of teacher performance (Manatt and Daniels, 1990).

1. Teacher and student attendance has a considerable impact on student achievement. Five to seven days' absence appears deleterious for students. A slightly longer period of absence for teachers, say seven to ten days, is significant. Attendance almost always seems to be an important factor in students' achievement in mathematics, less so in reading. Investigators conducting applied or theoretical research experiments simply must pay more attention to how much time intervention training causes teachers to spend away from students.
2. All subordinates want to "evaluate the boss." We found that measuring school climate affords a prime opportunity for teachers to give principals feedback on what teachers expect from them and what they perceive principals are giving them.
3. Perhaps the biggest surprise in the total-systems approach is the significant effect of pre- and posttesting (with proper reports to teachers) on student achievement. The advantage of twice-a-year testing with criterion-referenced measures is so great that the SIM research team stressed this link extensively in the next series of experiments.
4. Finally, how much does performance appraisal cost? Using 'ahead of the art" benefit-cost-analysis techniques, we found that teacher performance evaluations cost from $116 to $242 per teacher per year across four of the SIM school organizations. Even Breck School, which used a multiple evaluation team to determine merit pay, spent only $216 per teacher per year. Twice-a-year norm- and criterion-referenced testing cost $5 per pupil per year.

The Total-Systems SIM Model

Developing and field testing the model simultaneously in five very different school organizations enabled us to discover the power of outcomes, standards, and assessment-driven instruction almost a decade before such techniques became the vehicle for school transformation nationwide (Manatt, 1993). In the experimental analysis of the data gathered from 1980-84, three treatment groups were used. Each group's students were pre- and posttested with criterion-referenced tests. One group of teachers received massive staff development experiences centered on making them more effective teachers, one group had no training but their students were pre- and posttested. The third group had no training and their students were posttested only. To our surprise, pre- and posttesting had as positive an effect of all of the staff development combined. From that time forward, SIM projects all placed greater emphasis on curriculum renewal, curriculum alignment, and criterion-referenced assessment. Figure 3 illustrates the model in conceptual form.

Figure 3. School Improvement Model.



The completed model had eleven components.

1. LONG-RANGE STRATEGIC PLANNING
   - The belief systems of the school organization are reviewed.
   - The process of external and internal scanning is discussed.
   - Goals and objectives are written
   - Action plans are defined.

2. STAFF DEVELOPMENT
   - A committee within the school organization is charged with the responsibility of studying the research-based, staff development programs and selecting those which appear to be most appropriate for them.
   - Other criteria to be considered when selecting the program are the amount of time needed for instruction of the staff development program, any previous experience with the concept within the organization, and the cost.

3. CURRICULUM DEVELOPMENT
   - This component concentrates on helping to renew and/or define curriculum. The initial step is usually determining what changes are needed and then choosing the curricular area(s) on which to focus, writing philosophic statements, selecting the subject area program goals, and organizing the content.
   - The results of this process should be a curriculum guide or resource book which is used for making decisions about what is to be taught at which grade level and to what extent the learning should occur.

4. DEVELOPMENT OF CRITERION-REFERENCED MEASURES (CRMs)
   - This component focuses on writing/selecting assessments so they are embedded into what is taught.
   - Models of instruction are taught which focus on the commonly used paper-and-pencil assessments as well as alternative assessments.,
   - Criterion-Referenced Measures are designed to determine whether a student has achieved mastery of the learner outcomes which have been taught.

5. CURRICULUM ASSESSMENT
   - The basic premise upon which this component functions is that there needs to be a tie among the written/taught/tested curricula.
   - The process assists in providing direction so that a school organization behaves like a system rather than a group of parts.
   - A review of written documents along with on-site interviews are used to gather information about the basic premise.

6. THE COACHING PROCESS
   - The major purpose of this component is to teach skills found in a helping relationship.
   - The primary function is to assist others in developing "executive control" of a skill/concept/understanding as it is transferred to the teaching/learning situation.
   - Effective teaching behaviors are the focus of discussion as one learns to use the technical skills which are employed in this process.

7. PERFORMANCE EVALUATION SYSTEMS
   - Systems are developed for teachers, professional support staff, and administrators which are specific for the school organization. They are developed through the efforts of a representative group, the stakeholders' committee.
   - Input for each system is generated by discussing the structured activities and reaching consensus on the philosophic premises, performance criteria, and operational procedures.
   - These systems are piloted with a few people, revised by the stakeholders' committee, and then recommended to the Board of education for adoption. (Board evaluation is a discussion of "What is" and "What ought to be." The discussion is conducted by a facilitator for the Board of Education.,)

8. SUPERVISION/EVALUATION SKILLS
   - Skills-building sessions are taught for supervisors/evaluators so they know how to use the systems.
   - The skills included in these sessions are (a) analyzing instructional plans, (b) gathering data and analyzing them, (c) conducting conferences, (d) synthesizing data, (e) completing an end-of-cycle report, and (f) writing growth plans.

9. MANAGEMENT OF EVALUATION DATA (CATE/S)
   - This process can be used to generate reports by evaluator, district, and building.
   - Staff development needs can be identified by reviewing these reports.

10. SUPERVISING THE MARGINAL TEACHER/ADMINISTRATOR (Defining an Intensive Assistance Plan)
   - This component focuses on those teachers or administrators who are not meeting a district's standards. It is a helping routine that is a subsystem of the school organization's teacher, professional support staff, or administrator performance evaluation system.

- This process utilizes resources beyond that of a single evaluator because a team brings skills, knowledge, and time to the supervisory process.
- A sequential series of steps must be defined to help the employee improve his/her performance on a specific criterion.

11. CLIMATE
    - An instrument which analyzes administrative functions and climate has been developed. Teachers, building administrators, and central office personnel respond to items in each section of the instrument.
    - The six administrative functions are ranked for expectation and effectiveness on a scale of 1 to 5. The responses on the climate section are ranked on an eight-point scale.
    - After the data are analyzed, follow-up sessions are held. Suggestions as to "What can be done to improve?" are presented.

Next we had an opportunity to develop a longitudinal study of how the total systems approach to performance appraisal impacts teaching and learning. Hot Springs County District No. 1 in Thermopolis, Wyoming agreed to develop all of the SIM components in a five-year timeframe carefully monitoring school climate each year and student achievement as measured by the SRA tests. Climate stayed high and achievement increased dramatically (Manatt and Holzman, 1991). See Table 1.

Table 1. Student achievement by school--Percentile composite scores, national norms.[a]

	Year					
Level	Year One (1985)	Year Two (1986)	Year Three (1987)	Year Four (1988)	Year Five (1989)	Change
1 - 4	61	71	78	78	81	20
5 - 8	59	62	67	69	70	11
9 - 12	58	58	59	62	67	9
District Composite	59	65	69	71	73	14

[a]Science Research Associates Student Achievement Tests administered statewide in Wyoming.

Curriculum Assessment and Feedback in a Platform

Now let's move quickly over the past 10 years and examine SIM today. The SIM model has been developed with a series of school districts and consortia of school districts in Minnesota, Wyoming, Florida, Arizona, and Iowa since 1978. This is the fulltime responsibility of my SIM research partner, Dr. Shirley Stow.

Using mathematics as the example, the steps include:

1. Assessing the current status of content and curriculum goals for the target subject, mathematics.

2. Infusing life-long learning standards for mathematics into the instructional delivery system.
3. Renewing the present scope and sequence of the district's mathematics program.
4. Bridging from the existing and renewed scope and sequence of the district to learner outcomes and assessment materials based on research by the SIM team at Iowa State University.
5. Enabling the district's teachers and administrators to use the new goals, standards, and assessment materials.
6. Computerizing all curriculum management activities and the assessments for mathematics.
7. Fieldtesting all aspects of the goals/standards/assessment model for mathematics. (Pilot test with at least 100 students per grade per subject.)
8. Assessing the effectiveness of the model and making revisions.
9. Training all district stakeholders to operate the model and to use the records generated. (Manatt, 1994)

Two of our research partner districts will be used as examples of robust improvement of teaching and learning using the SIM approach: Monroe County, Florida (the Keys) and Gilbert Community Schools in Iowa. In Monroe County, under the leadership of Robert Walker, superintendent, and Phyllis Allen, director of curriculum, five years of development and experimentation has resulted in pre- and posttesting of all subjects, all grade levels. A gain score is computed for each child, class, section, grade, school, and district. Robert Walker and Phyllis Allen know more about achievement of their students than any management team in the country. All of the CRT test analyses are completed by mainframe computer at the Durham Computer Center at ISU.

In a much smaller district, Gilbert, we have taken the formative testing approach with all curriculum management, test generation, and scoring by the National Computer System (NCS) platform called Performance Plus©. Doug Williams, superintendent, and David Ashby, elementary principal, have a complete diagnosing, assessing, and reporting system that serves teachers, students, parents, and the district.

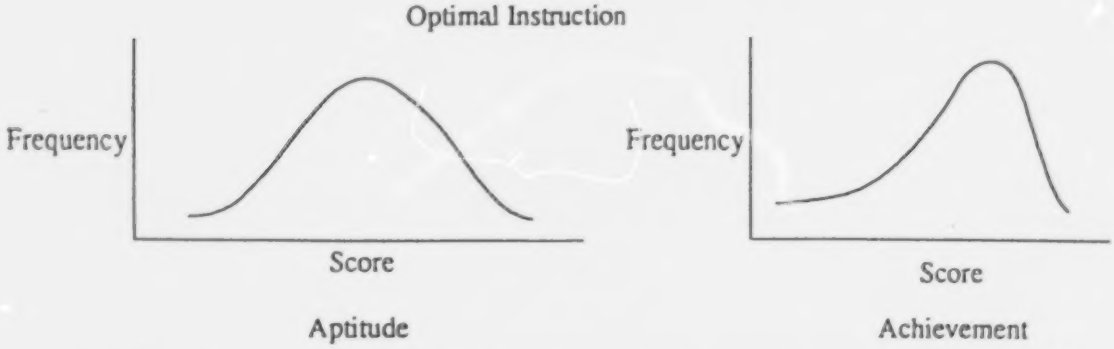
The foundation for much of this work was the pioneering research done by Benjamin S. Bloom in the mid '60s, with his studies of human variabilities, at the University of Chicago. He became interested in how teachers might change teaching and learning practices in order to provide higher quality instruction for a larger portion of the students taught. Bloom was interested in discovering ways where all students would learn well and, as a result, "reduce the variability that typically exists in students' level of achievement" (Guskey, 1988, p. 48).

When instruction was uniform, with tests and quizzes given as a summative evaluation, the amount of achievement is fairly dependent on student aptitude. Bloom discovered that in most cases only about 20 to 30 percent of the students in a class really learn well what the teacher sets out to teach (Guskey, 1988).

Although Bloom (1981) believed that dividing material into small sequential units and checking students' learning by means of quizzes and tests at the end of each unit was a useful instructional technique, he felt that the tests and quizzes typically used by teachers did little more than tell them who was doing well and who was not. Missing was some form of more meaningful feedback and corrective process. Such checks on student performance could be used not only for grading and evaluation but also to diagnose individual learning difficulties (feedback) and to prescribe specific remediation or reteaching procedures (correctives.)

As a result of using a group based master learning[b] Bloom (1981) firmly believed that this sequence of formative testing and systematic correction of learning difficulties would provide each student with a better and more appropriate quality of instruction than is possible with more traditional approaches. He believed that students would become more similar in terms of their achievement, their motivation for future learning, and "perhaps even the rate at which they learn"(Guskey, 1988, p. 52). Figure 4 illustrates the relationship of aptitude to achievement under what Bloom described as "optimal instruction."

Figure 4. The relationship of aptitude to achievement under optimal instruction



If students were normally distributed on aptitude but each received optimal (quality) instruction and received the learning time needed, then the majority of the students would attain mastery. There was little correlation between aptitude and achievement under these conditions. With the same level or standards of achievement expected, under uniform instruction 80 percent or more of the students in a class would typically achieve what 20 to 30 percent did under more traditional instruction (Bloom, 1981).

Our SIM system uses maximum feedback to teachers, students, parents, and administrators. We are not forcing mastery teaching; however, teachers teach differently and students learn differently when all parties have a continuous feedback loop with a rich array of information.

Our conclusions to date are that a microcomputer-based system (1) is a lot of work to set up, (2) puts testing to work for the teacher, (3) causes teachers to become more reflective, (4) makes very user friendly reports of progress to students and parents, and (5) provides for the key elements of the GOALS 2000 Educate America Act and its requisite Opportunity to Learn measures.

---

[b] There are three primary forms of mastery learning. The Personalized System of Instruction (PSI) is used primarily at the post-secondary level. A second form is called Continuous Progress (CP) mastery learning theorized by Cohen. Group-based mastery learning or Learning for Mastery (LFM) is the third form (Bloom, 1981). This third form is the primary focus of our efforts.

References

Anderson, C.J., A.,S. Barr, and M.,G. Bush. 1925. *Visiting the Teacher at Work.* New York: Appleton-Century.

Bloom, B.S. 1981. *All Our Children Learning: A Primer for Parents, Teachers, and Other Educators.* San Francisco, CA: McGraw-Hill.

Cogan, M.L. 1973. *Clinical Supervision.* Boston: Houghton Mifflin.

Duke, D.L. and R.J. Stiggins., 1986. *Teacher Evaluation: Five Keys to Growth.* Washington, DC: Joint publication of American Association of School Administrators, National Association of Elementary School Principals, National Association of Secondary School Principals, and National Education Association (NEA Library).

Goldhammer, R. 1969. *Clinical Supervision: Special Methods for the Supervision of Teachers.* New York: Holt, Rinehart, & Winston.

Guskey, T.R, 1988. *Improving Student Learning in College Classrooms.* Springfield, IL: Charles C. Tomas.

Hidelbaugh, E.J. 1973. *A Model for Developing A Teacher Performance Evaluation System: A Multiple Appraiser Approach.* Unpublished Ph.D. dissertation, Iowa State University, Ames, IA.

Manatt, R.P. 1993. "The Changing Paradigm of Outcomes and Assessment," *International Journal of Educational Reform*, 2(1): 86-95.

Manatt, R.P. 1994. "Improving Learning vs. Auditing What Has Occurred," an address to the Upper Midwest Assessment Conference, NCS, Eden Prairie, MN.

Manatt, R.P. and B. Daniels. 1990. "Relationships Between Principal Ratings of Teacher Performance and Student Achievement," *Journal of Personnel Evaluation in Education*, 4:189-201.

Manatt, R.P. and G.R, Holzman. 1991. "Criterion-Referenced testing for Outcome-Based Education," *Educational Considerations*, 18(2):22-24.

Manatt, R.P. and S.B. Stow. 1984. *Clinical Manual for Teacher Performance Evaluation.* Iowa State University Research Foundation, Ames, IA.

Manatt, R.P. and S.B Stow. 1986. "Developing and Testing A Model for Measuring and Improving Educational Outcomes of K-12 Schools," Technical Report. Ames: Iowa State University.

Manatt, R.P., K. Palmer, and E. Hidlebaugh. 1976. "Evaluating Teacher Performance With Improved Rating Scales," *NASSP Bulletin*, 60:21-24.

Medley, D.M. 1992. "Teacher Evaluation," in *Encyclopedia of Educational Research*, M. Alkin, ed. 4: 1345-1352.

Medley, D.M. and H. Coker. 1987. "The Accuracy of Principals' Judgments of Teacher Performance,"*Journal of Educational Research*, 80(4): 242-247.

Scriven, M. 1988. "Duty-Based Teacher Evaluation." *Journal of Personnel Evaluation in Education*, 1: 319-334.

Scriven, M., P. Wheeler, and G. Haertel. 1991. *Teacher Evaluation Models Project*. The Center for Research in Educational Accountability and Teacher Evaluation. Western Michigan University, Kalamazoo, MI, p. 7.

Stow, S.B. and J.E. Sweeney., 1981. "Developing A Teacher Performance Evaluation System," *Educational Leadership*, 38:538-541.

jmm (7/94)